

Model Differences

Generating Semantically Relevant Difference Scripts.

Ulf Schwekendiek

Modelling and Simulation Based Design 2008
Winter Term COMP 763B
McGill University, Montral, Canada
Student ID: 260225272
ulf.schwekendiek@gmail.com

1 Introduction

Detection computation of differences in information models are essential to many development and management practices. In most of the cases a differences is defines as a script explaining how to get from a source information to a target one. The GNU diff tool for instance is highly used in current version control systems such as CVS or SVN in order to compute such scripts. This tool is tailored for detecting plain text source code difference in a very efficient way. It uses the longest common subsequence algorithm in order to compute an upgrade actions such as add, delete and upgrade.

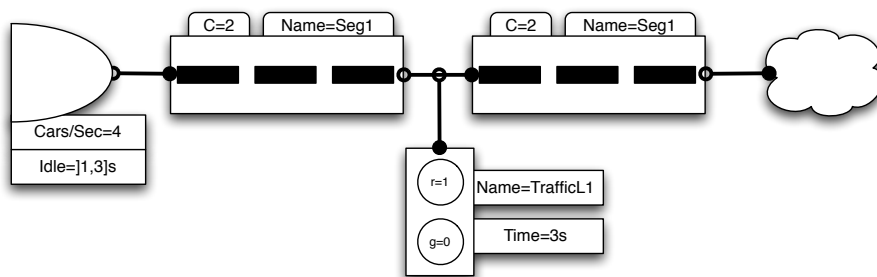


Fig. 1. Sample traffic net model.

Sadly this tool leads in Model-Driven Engineering (MDE) approaches to very unsatisfactory results. In this approach I use the possibility that every model

can be mapped to its XML representation. The GNU diff tool can take now the flatten textual input and compare two models through their XML representation. However this leads to very bad results. There is another algorithm called X-Diff which yields to much better results and should rather be used. This algorithm computes a nearly optimal update script in most of the cases as long as the changes between the source and target model are not that big. However when we are creating this XML based model information we generate them usually with some tools. This tools store in many cases a lot of noise in the model's XML representation. For instance in a simple traffic net model like in Figure 1 not only the model relevant data is stored in XML. Last change time and position of each model element are also stored in many case tools. But if we compare Figure 1 and Figure 2 the only change really want to be aware of would be a change at the semantic level. Attributes are changing and a traffic light is gone. Therefore the meaning of this model changed. For a version control management system this information should not be relevant and hence be suppressed.

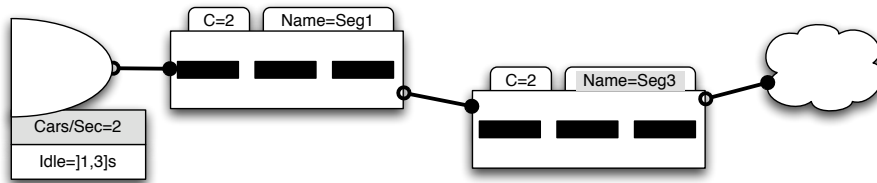


Fig. 2. Changed sample traffic net model from Figure 1. Only the grey colored attributes, and the absence of the traffic light are important.

Based on the X-Diff algorithm I will refine the way differences for models will be computed and represented. I will add a set of rules for the algorithm in order to separate semantic relevant information from irrelevant ones. I will also define a prototype of this work and provide time measurements and other relevant tests.